SAGE III Quality Assurance Plan

Draft

March 1997

Science Applications International Corporation
Hampton, Virginia

SAGE III Quality Assurance Plan

March 1997

Prepared Under Contract NAS1-19570

Prepared for:
NASA Langley Research Center
Hampton, VA 23681

Prepared by:

Robert E.Veiga	Date

SAGE III Algorithm Development Team, SAIC

Prepared by:

Anne C. Edwards	Date

SAGE III Data Modeler, SAIC

Reviewed by:

Mary T. Osborn	Date

SAGE III/SAIC Task Leader, SAIC

Approved by:

William P. Chu	Date

SAGE III Science Manager, NASA/LaRC

Science Applications International Corporation

1 Enterprise Parkway

Hampton, Virginia 23666

Table of Contents

1. Introduction

1.1. Identification

In response to the Earth Observing System (EOS) Data Quality Panel, the Stratospheric Aerosol and Gas Experiment III (SAGE III) Quality Assurance (QA) Plan presents a collection of procedures which will be applied to the SAGE III telemetry and data products during operational EOS Data and Information System (EOSDIS) Core System (ECS) processing, and in post-processing at the NASA Langley Research Center (LaRC) SAGE III Science Computing Facility (SCF), in order to ensure that the archived data granule's content meets expected standards of accuracy, and is flagged in instances where standards are not met. These procedures are needed to assure a high quality SAGE III data product set, and the QA information inserted into the data products will aid a wide range of users who access these data. The QA process is intended to clearly indicate the usability a particular data product may have in a specific application.

The QA Plan defined in this document also includes support for the areas identified by the Data Quality Panel related to instrument calibration, monitoring, and data validation.

1.2. Scope

All QA products defined in this plan will encompass SAGE III Level 1B (L1B) and Level 2 (L2) data products, and all QA analysis will take place solely within the Product Generation Executive (PGE) and in the Science Computing Facility (SCF). No Distributed Active Archive Center (DAAC) analysis is required for QA functions. It is critical to accurately identify the operations of data extraction and analysis that will occur in the L1B/L2 code in a timely fashion because of the tight schedule leading to Version 2 code delivery and subsequent infrequent opportunities for change.

A brief overview of the SAGE III system is presented, followed by the description of the central elements of the QA Plan. There are five areas comprising the SAGE III QA Plan: QA data; QA metadata; engineering telemetry analysis; trending and consistency testing; and analysis tools. These activities encompassing these areas are described in the following sections.

1.3. Experiment Description, Objectives, and Data Products

SAGE III measures the attenuation of solar or lunar radiation by the Earth's atmosphere due to scattering and absorption by atmospheric constituents during each sunrise/moonrise and sunset/moonset encountered by its spaceborne platform. The moonrise and moonset measurements are made when the atmosphere is not directly illuminated by the Sun. SAGE III consists of the following three subsystems: pointing, imaging, and spectrometer subsystems. The pointing subsystem consists of a scan mirror which acquires the radiant target, and performs elevation scanning across the target. A measurement is considered to occur at the point along the line of sight from the instrument to the target at which it comes closest to the Earth's surface. The altitude of that point above the Earth's surface is commonly referred to as the tangent altitude. The imaging subsystem produces a focused image of the target at a focal plane where the "science" aperture, that defines the instrument's instantaneous field of view (0.5 km at the tangent altitude), is situated. The spectrometer subsystem is situated behind the science aperture and consists of an 809-element CCD to detect solar radiation at 70 to 80 wavelengths from 280 to 1040 nm with 1 to 2 nm spectral resolution. An additional photodetector is used to measure radiation at 1550 nm. The instrument is scheduled to launch in August 1998 on a Russian METEOR-3M spacecraft.

The SAGE III retrieval algorithm is the procedure that converts measured instrument counts into vertical profiles of the molecular density of gaseous species, aerosol extinction at 8 wavelengths, temperature, pressure, and clouds. The science objectives to be accomplished by SAGE III are the following:

1. Characterize tropospheric as well as stratospheric clouds and investigate their effects on the Earth's environment, including radiative, microphysical, and chemical interactions
2. Determine long-term trends in gaseous species and temperature
3. Provide atmospheric data essential for the interpretation and calibration of other satellite sensors, including EOS instruments
4. Investigate the spatial and temporal variability of these species in order to determine their role in climate processes, biogeochemical cycles, and the hydrological cycle

At launch, the SAGE III data products consist of a Level-1B transmittance product, and numerous Level-2 products consisting of the gas, aerosol, temperature, and pressure profiles, as well as a cloud detection product. No data products beyond Level-2 are currently planned.

1.4. SAGE III SDP and QA Overview

The DAAC or the SCF hosts execution of the various PGEs which generate the species products and transmission. At run time some QA functions are performed within the PGEs and others are performed by the QA Executive (QAE). All Data Products, QA and Production Logs are archived at the DAAC, and also shipped to the SCF for further QA analysis. The schematic of the data flow, shown in Fig. 1, concentrates on delineating the essential elements of the QA process. Instrument data needed by the PGEs comes from the Wallops Flight Facility (WFF) through Mission Operations Control (MOC), and the Data Assimilation Office (DAO). MOC processing generates orbit state vectors, accompanied by low-level QA on both the Level-0 (L0) data and the definitive orbit. All instrument telemetry limit checking is also performed within the MOC modules. Meteorological input data from the DAO characterizes the state of the atmosphere at the SAGE III sampling location, and a low level QA check for integrity is performed. All input data QA is pushed to SCF for assimilation.

The PGE requires specific inputs, and internal QA is performed and routed through the QAE to the DAAC and the SCF where extended QA is performed at the SCF. QA processing feeds back to the PGEs through the PGE input files.

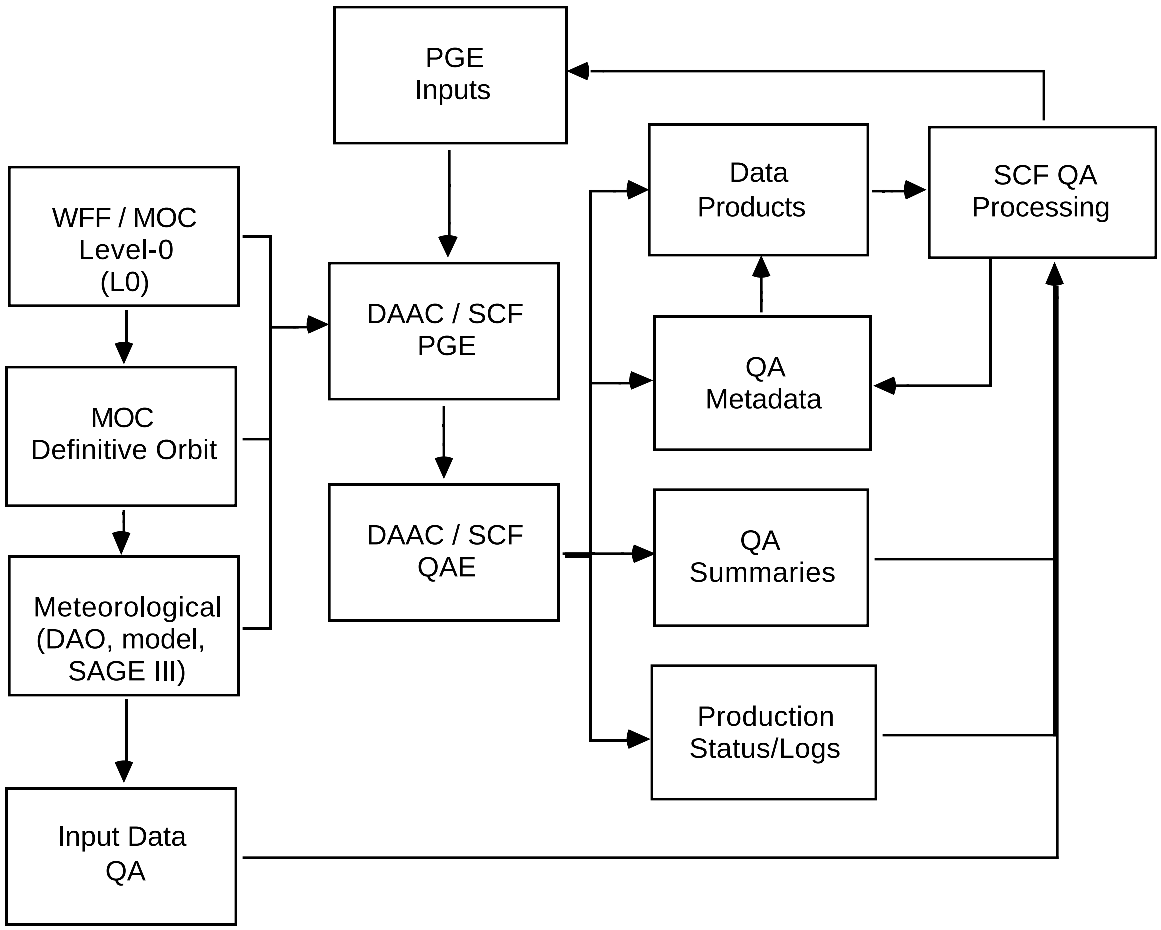

Figure 1. SAGE III Data Flow

1.4.1. Granule QA Data

During operational SAGE III L1B and L2 processing, a series of tests are conducted to assess the data product quality or characteristics at the altitude level (data plane) in the granule profiles. Some of these values indicate serious problems with the resultant data, and some indicate conditions that are not known to corrupt the results but are nonstandard in some fashion. Details can be found in section 2.

1.4.2. Granule QA Metadata

These metadata include the EOS core and non-core attributes, and represent results of evaluating the whole granule, or individual profiles within the granule. The values for these QA items are of string type, and are meaningful for all data users. These values are assigned either at run time by the QAE, or assigned as a result of SCF QA processing. Details described in section 3.

1.4.3. Engineering Data

The SAGE III telemetry stream consists of measurements of internal instrument state. These include temperatures, voltages, times, currents, and other instrument parameters. These values will be monitored for values exceeding nominal limits and statistical deviations from expected patterns. Basic instrument health and safety monitoring is controlled by the MOC. Additional engineering telemetry analysis is performed and assimilated in the SCF QA processing. The QA tests and collected statistics are described in section 4.

1.4.4. Trending and Consistency Analysis

The SAGE III instrument and project are a synergistic system aimed at satisfying the science objectives above, and for this purpose the instrument design includes built-in redundancy for guaranteeing stability. Multiple techniques might exist to determine the same or related quantities, and we choose the most reliable for monitoring instrument or data product stability. The other techniques can be evaluated for consistency. Nominally stable behavior can be tested for temporal changes. These tests provide a level of protection against large classes of errors. The trending and analysis tests are described in section 5.

1.4.5. Analysis Tools

Specialized analysis tools are needed to perform the tests and make the reports described in this document. Some exist now, some can be purchased off the shelf, and some tools will need to be built. Section 6 describes some of those tools and how we may obtain them.

1.5. Parent Documents

1. Mission to Planet Earth Strategic Enterprise Plan 1996-2002, NASA HQ MTPE, May 1996.
2. Execution Phase Project Plan for Earth Observing System (EOS), GSFC 170-01-01, Rev. A., May 1995.
3. Science Data Processing Segment (SDPS) Database Design and Database Schema Specifications for the ECS Project, 311-CD-002-005 EOSDIS Core System Project, May 1996

1.6. Supporting Documents

1. SAGE III Algorithm Theoretical Basis Document: Transmission Data Products, LaRC XXX-XXX-XXX, November 1996.
2. SAGE III Algorithm Theoretical Basis Document: Temperature and Pressure Data Products, LaRC XXX-XXX-XXX, November 1996.
3. SAGE III Algorithm Theoretical Basis Document: Aerosol Data Products, LaRC XXX-XXX-XXX, November 1996.

4. SAGE III Algorithm Theoretical Basis Document: Nitrogen Dioxide Data Products, LaRC XXX-XXX-XXX, November 1996.
5. SAGE III Algorithm Theoretical Basis Document: Nitrogen Trioxide Data Products, LaRC XXX-XXX-XXX, November 1996.
6. SAGE III Algorithm Theoretical Basis Document: Chlorine Dioxide Data Products, LaRC XXX-XXX-XXX, November 1996.
7. SAGE III Algorithm Theoretical Basis Document: Water Vapor Data Products, LaRC XXX-XXX-XXX, November 1996.
8. SAGE III Algorithm Theoretical Basis Document: Cloud Presence Data Products, LaRC XXX-XXX-XXX, November 1996.
9. SAGE III Algorithm Theoretical Basis Document: Ozone Data Products, LaRC XXX-XXX-XXX, November 1996
10. Software Requirements for the Stratospheric Aerosol and Gas Experiment III (SAGE III) (SAGE III Transmission), LaRC

1.7. Acronym List

ATBD	Algorithm Theoretical Basis Document
A/D	Analog to Digital
A&E	Activation and Evaluation
Ball	Ball Aerospace Corporation
CSCI	
DAAC	Distributed Active Archive Center
DAO	Data Assimilation Office
DC	Direct Current
DN	Digital Number
ECS	EOS Core System
EOS	Earth Observing System
IT	Instrument Team
LaRC	Langley Research Center
L1B	Level 1B
L2	Level 2
MLR	Multiple Linear Regression
MOC	Missions Operation Control
PGE	Product Generation Executive
PSC	Polar Stratospheric Cloud
SAGE III	Stratospheric Aerosol and Gas Experiment III
SAIC	Science Applications International Corporation
SCF	Science Computing Facility
SDP	Science Data Production
S/C	Spacecraft
RSS	Residual Sum of Squares
RVDT	Rotating Variable Differential Transformer
WFF	Wallops Flight Facility

2. QA Data

2.1. Data Plane

A data plane will accompany each species profile. This plane is defined as a set of QA profiles which correspond, in altitude, to the data product profiles. Some possible values are described in Table 1, and future data planes are being considered.

Table 1 Data Plane Accompanying Each Profile

#	Data Plane Descriptor	Value	Level
1.	Probability that a profile value was sampled within the circumpolar vortex	0.0 - 1.0	L2
2.	Statistically anomalous profile value with respect to an ensemble of profiles for that altitude.	0.0 or 1.0	L2
3.	Possible auroral contamination for lunar species retrievals.	0.0 or 1.0	L1B, L2
4.	Spatial inhomogeneity in the slant path	0.0 or 1.0	L1B

2.2. Fill Data

These are profile-specific indicators that a retrieval could not be completed successfully. Since the SAGE III profiles are altitude dependent, the fill-values are placed in the data product profiles at each altitude where problems may have occurred. Level 2 processing software must check for these fill values in the L1B transmissions even if no QA flags have been set for the corresponding L1B profile. Special care should be used to avoid averages, smoothing, or statistical operations that would use these fill data as valid geophysical parameters. The fill values are defined at run time as floating point numbers and are listed in Table 2 (specific values are TBD).

Table 2 Fill Data

#	Fill Data Descriptor	Fill Value	Level
1.	Signal level low (no inversion	TBD	L1B

	possible)		
2.	Dead detector(s) in a pixel group	TBD	L1B
3.	L1B transmission failure	TBD	L1B
4.	L2 transmission failure	TBD	L2
5.	Saturated detector in a pixel group	TBD	L1B

3. QA Metadata

3.1. Core QA Metadata

Each SAGE III L1B or L2 granule will incorporate mandatory QA information derived from tests made during the processing of that granule. EOS defines the following QA metadata attributes:

3.1.1. QACollectionStats

The *AutomaticQualityFlag* (AQF) and *ScienceQualityFlag* (SQF) will be assigned the character string values of 'pass' or 'fail'. The value represents a distillation of information collected by the QAE at run time. The SQF is set after a more in-depth SCF analysis. These metadata would express various problems/conditions of the data/instrument that would have meaning to the data user and PGEs. *QualityFlagExplanation* will point the user to a document that explains the two quality flags in depth. *OperationalQualityFlag* is optional and will not be used.

3.1.2. QAStats

Each SAGE III L1B granule will contain 85 data channels, each having its own nominal altitude coverage as described in the Transmission ATBD. Each SAGE III L2 data granule will contain either 20 (solar) or 13 (lunar) species profiles. Each species profile has its own nominal altitude range as described in the species ATBDs. The QAStats attributes for each granule will be defined as:

1. *QAPercentMissingData*. Average within the granule of the percentages of each channel/species profile containing fill values.
2. *QAPercentOutofBoundsData*. Average within the granule of the percentages of each wavelength/species profile containing values

exceeding the maximum or minimum reasonable bounds for that wavelength/species.
3. *QAPercentInterpolatedData*. Always 0% for Level 1B and L2 products.

3.1.3. Accessibility Code

The proposed core metadata attribute "accessibility code" would indicate that the granule should be held for inspection by the DAAC or SCF, the granule could be made accessible to the general public, the granule could be made accessible to other ITs only, or some other condition. This attribute has not been completely defined or included in the data dictionary as yet.

3.2. Non-Core QA Metadata

Each SAGE III L1B and L2 granule will incorporate summary information derived from tests made during the processing of that granule. The QAE uses a function of a subset of these values to set the AQF. Some will apply to the individual profiles; the remainder will apply to the entire granule. Potential candidates for these metadata are described in Table 3.

Table 3. Summary QA Parameters

#	Parameter Name	Description	Impact
1.	Non-Routine Instrument Operations	Indicates that a non-routine event is occurring which may or may not affect the retrieval. Examples of such events are: one or more critical engineering variables out of limits;	Calibration product usable, but may exceed budget and specification.
2.	Marginal Calibration Data	Indicates some missing calibration data which could impact the calibration error budget, and may or may not exceed the SAGE III specification.	Calibration product usable, but may exceed budget and specification.
3.	Bad Calibration Data	Indicates missing calibration data and/or relevant engineering data to either provide no retrieved product or a product which exceeds the calibration budget and specification.	No usable data product.

4.	A & E Phase	Indicates that the spacecraft and instrument are in the Activation and Evaluation stage of the mission.	Calibration may not meet specification because on-orbit characterization values still need to be determined based on an analysis of the data from this phase. All uncertainty estimates are set to 100% of their respective units.
5.	Tropopause altitude	Altitude of the tropopause derived from the pressure-temperature algorithm (km)	None
6.	Systematic Uncertainties	Systematic uncertainties (1σ) in each species profile. These values are determined prelaunch or through validation.	Needed for users if assessment of the total uncertainty.
7.	Noisy Detectors	List of the CCD detectors known to display noise beyond the specification SNR.	Depending on the characteristics of the noise, the data product uncertainty estimates at a given time may be too low.
8.	ZigZag test	Species exceeds reasonable bounds as established from climatological profiles in the PGE input files.	Unusable profile
9.	Model meteorological only	Model meteorological model profile was used as the first guess in the inversion rather than measured profile.	Needed for QA on the data product p/T retrieval.
10.	SR/SS mixing	Satellite event opposite of S/C event.	Valuable for data users studying diurnal effects.
11.	Profile truncates in stratosphere	Profile is truncated at higher altitude than normal.	Loss of data at lower altitudes or profile is suspect.

12.	Aurora detected	Auroral spectra detected during lunar event.	Event is suspect at altitudes corresponding to the auroral emissions.
13.	Convergence test	Species retrieval did not convergence after TBD steps.	Increased uncertainty in species profile.
14.	Onion peeling algorithm	Version of the onion peeling algorithm used to derive certain species.	Internal QA
15.	Ozone algorithm	Version of the ozone retrieval.	Internal QA

4. Engineering Data

4.1. By granule

For each of the engineering data variables within each granule the following quantities are calculated:

N, Number of observations
Mean, Mean value of the observations,
Sigma, Standard deviation of the observations,
Min, Minimum value recorded in the granule,
Max, Maximum value recorded in the granule ,

These results will be transferred to the SCF at the end of each L1B within granule process for subsequent time series analysis. The intent is to build a complete record of these engineering parameter statistics over the life of each SAGE III instrument. We expect these records to be valuable for studies of the time dependent state of the SAGE III instrument and correlative studies with retrieved parameters.

4.2. Within the SCF

Within the SCF several analyses will use this engineering data. One of the first activities during A&E will be forming histograms of raw DNs through each of the A/D converters used for engineering data. DN values, or patterns of values that never occur will be analyzed and traced in order to discover probable causes of the anomalous values. Broken wires or internal D/A flaws could create such patterns. Time series analysis tools in the SCF plot the data extracted from the L1B granules to compare the values relative to high and low limits will be needed. Trending and extrapolation of trends will also be needed with polynomial and harmonic fitting functions.

5. Trending and Consistency Analysis

5.1. Introduction

It is critical to accurately identify the operations of data extraction and statistical analysis that will occur in the L1B/L2 code in a timely fashion because of the tight schedule leading to Version 2 code delivery and subsequent infrequent opportunities for change. The information in Table 4 provides the priorities for the activities listed in the SAGE III QA Plan and where processing is performed.

Priority 1 means that changes are required in the L1B/L2 CSCIs. Full details must be specified for inclusion in version 2.0 by late 1997.

Priority 2 means that the test will take place in the SCF and is required at launch. The algorithm and needed software should be coded, tested, and under configuration control by launch minus one year, June 1997.

Priority 3 means the analysis will take place in the SCF and will be completed as resources permit.

An X in a column means that some of the unique processing for the named QA test will take place in this processing system.

Table 4 Priorities and Processing Locations

QA Test	QAE	SCF	Priority
QA data			1
QA metadata			1
Engineering trend data			1
1. Solar Calibration Descriptive Statistics			
2. Wavelength Calibration			
3. Mirror Calibration			
4. Mirror Rate			
5. Intra-orbit variations			
6. Trapped Radiation Effects in the CCD DC Offset			
7. Coherent Noise in the Normalized Solar Scans			
8. Temperature Telemetry Change Point Detection			

9. Auroral Emission in Lunar Measurements			
10. Lunar Disk Registration			
11. Spatial Homogeneity Test			
12. Convergence Test			
13. MLR Wavelength Shift			
14. Dual Inversion Ozone Agreement			
15. Cloud Product			

Any trending or consistency analysis described in section 5 follows a format for answering the following questions:

1. Name for the analysis
2. SAGE III data being analyzed
3. Characteristic of SAGE III, or assumption used in an algorithm, to be verified or studied in the analysis
4. Method of analysis
5. Schedule
6. Format for presentation of the analysis results
7. Special tools needed to perform the analysis

A section may have several answers. An analysis could lead to an article in a peer reviewed journal, a presentation at a SAGE III Science Team Meeting, a data set generated as an HDF file and archived at the DAAC, and a request to the SAGE III Team Leader for a change in the retrieval algorithm.

The analyses described below are intended to verify the assumptions that have been made in the process of species retrieval from SAGE III measurements. In general, the expected results will only be of interest to the algorithm developers, and thus reports of most results will be made to the SAGE III Algorithm Group. Failures of the retrieval algorithm discovered during these analysis will be reported to the SAGE III Team Leader by the existing Algorithm Group.

5.2. Solar Calibration Descriptive Statistics

SAGE III data analyzed

SAGE III DN during solar calibration mode for the altitudes of 100 to 150 km.

Assumption

The SAGE III system will use as a radiometric calibration the exoatmospheric scans across the solar disk. Subsequent scans of the disk with the tangent altitude in the atmosphere will be normalized relative to the radiometric calibration scan. Various statistics resulting from the calibration scan impact the accuracy and precision of the retrieved solar transmission profile (output of L1B), and will be trended in SCF analysis.

Analysis

The SAGE III system requires solar scans to set the baseline for atmospheric transmittance calculation. After A&E the radiometric performance will decrease because of probable degradation of optical and electronic components, and thus cause a drop in the system's responsivity. During the solar calibration, numerous scans across the solar disk (up/down scans) are averaged within the L1B algorithm to produce the reference up/down scans. Sunspot counts are recorded and removed from the solar disk scans. At each scan the maximum DN, median DN, and standard deviation DN are analyzed for all detectors for both up and down scans. The time series of these statistics are transformed into the current best estimates of the channel degradation, and SNR for each detector.

The background subtraction DN depends on the part of the scan taken while the mirror direction turns from dark space back onto the solar disk. Time series of the mean background standard deviation DN are recorded.

The solar edge detection algorithm in L1B dictates the fraction of the solar disk used during the transmission retrieval. As the scan mirror sweeps onto or off the disk, the change in DN is measured by the standard deviation of the DN. This standard deviation is compared against an input threshold, and if it is less than the threshold value, then the solar disk is assumed to be contained in the FOV, and the data are used for transmission retrieval. Time series of the solar disk scan standard deviation will aid in determining the effectiveness of the edge detection algorithm in L1B.

Schedule

The solar scan data to perform this analysis will be collected every granule with solar calibration data. The analysis leading to a graph of the statistics described above will be done at least every months throughout the SAGE III mission.

Presentation and results

The SAGE III solar scan statistics will be saved in designated files for consistency tests, and made available at each meeting of the SAGE III Algorithm Group.

Special tools

The time series analysis tools will be needed.

5.3. Wavelength Calibration

SAGE III data analyzed

The data analyzed includes the SAGE III DN during wavelength calibration mode between the altitudes of 150 to 180 km.

Assumption

The SAGE III system will use the data from exoatmospheric scans across the solar disk to effect a wavelength calibration of the CCD. These data are sampled from the 809 pixels of the CCD at an effective rate of 6.4 Hz. The wavelength registration of the CCD is a map from pixel number to wavelength, and errors in this mapping will cause a systematic error to be propagated into the species retrieval. The uncertainty in the pixel-to-wavelength map are caused by optical degradation, CCD noise and temperature effects, as well as uncertainty in the reference solar spectrum.

Analysis

The SAGE III CCD is mounted on a copper block for thermal stability. However it is expected that temperature changes throughout an event may cause a slight distortion in the CCD, as well changes in the physical state of the grating and optics. The wavelength calibration algorithm in L1B calculates the best pixel-to-wavelength map based on sampling regions of the solar spectrum containing the deep Fraunhoffer lines. Several distortion parameters and their resulting variances result from the fit of the CCD response to these lines. Time series of the distortion parameters are analyzed for trends and stability of variance. Results are input to the PGE input and the QA metadata.

Schedule

The solar scan data to perform this analysis will be collected every granule with wavelength calibration data. The analysis leading to a graph of the statistics described above will be done at least every months throughout the SAGE III mission.

Presentation and results

The SAGE III wavelength calibration statistics will be saved in designated files for consistency tests, and made available at each meeting of the SAGE III Algorithm Group.

Special tools

The time series analysis tools, and a reference solar irradiance spectrum will be needed.

5.4. Mirror Calibration

SAGE III data analyzed

The data analyzed includes the SAGE III DN during mirror calibration mode between the altitudes of 180 to 300 km.

Assumption

The SAGE III system will use the data from exoatmospheric scans across the solar disk to effect a calibration of the system response to scan mirror angle. These data are sampled from the current pixels of the CCD table at the effective rate of 64 Hz. Uncompensated mirror angle response will cause a systematic error to be propagated into the species retrieval. The uncertainty in the mirror angle response are caused by optical degradation.

Analysis

Time series of the following quantities are trended for each wavelength in the CCD table:

1. polynomial coefficients of the fit of mirror angle vs. DN
2. coefficient standard errors
3. prediction error (extrapolation into angles for the atmosphere)

Schedule

The mirror data to perform this analysis will be collected every granule with wavelength calibration data. The analysis leading to a graph of the statistics described above will be done at least every months throughout the SAGE III mission.

Presentation and results

The SAGE III mirror calibration statistics will be saved in designated files for consistency tests, and made available at each meeting of the SAGE III Algorithm Group.

Special tools

The time series analysis tools needed.

5.5. Mirror Rate

SAGE III data analyzed

The data analyzed includes the SAGE III DN during an entire event from 300 km to the profile truncation point.

Assumption

The SAGE III elevation scan mirror motor sweeps the solar disk across the input aperture at a constant rate during one scan. Several effects, such as temperature changes, may cause perturbations on this assumed constant rate. The instrument manufacturer provides temperature dependent corrections in order derive the best estimate for the actual position of the scan mirror on the solar disk. The scan mirror position must be known accurately for subtangent point geolocation. The L1B algorithm uses a linear fit to the elevation scan mirror DNs to obtain the 'best' predictor of the mirror position.

Analysis

Time series of the following quantities are trended for all scans on each event:

1. R^2 of the linear fit to each scan
2. mirror motor (RVDT) temperature

3. number of outliers per scan

Schedule

The mirror scan data to perform this analysis will be collected every granule for all scans of each event. The analysis leading to a graph of the statistics described above will be done at least every months throughout the SAGE III mission.

Presentation and results

The SAGE III mirror scan statistics will be saved in designated files for consistency tests, and made available at each meeting of the SAGE III Algorithm Group.

Special tools

The time series analysis tools needed.

5.6. Intra-orbit variations

SAGE III data analyzed

A subset of the high rate (64 Hz) and low rate (0.1 Hz) engineering telemetry (IC subcom 1 and 2)

Assumption

As experienced with SAGE II, a thermal shock on the instrument as the platform came out of the umbra during a sunrise event affected the retrieval of certain gas species. The SAGE III performance change within an orbit can be monitored by the data from selected engineering outputs. It is especially useful to detect diurnal temperature changes since the CCD and imaging optics may be sensitive to both these effects.

Analysis

Schedule

Presentation and results

These results will be regularly reported and presented to the SAGE III Algorithm Group.

Special tools

The time series analysis tools will be needed.

5.7. Trapped Radiation Effects in the CCD DC Offset

SAGE III Data Analyzed

Monthly SAGE III CCD DC for low rate (0.1 Hz) and high rate sampling with no outlier rejection for all orbits in one day. Latitude, longitude, and time of the subsatellite point at the beginning of each granule.

Assumptions Tested

The masked detectors on the CCD array are used as the zero radiance source for the SAGE III spectral bands. The Ball design of all the SAGE III detectors required a nominal level of isolation from the trapped radiation environments encountered in a typical orbit at 1000 km. However there is a potential for charged particles to strike the CCD detectors producing electron-hole pairs which will contribute to anomalous output at the respective A/Ds. Thus the DC offset DNs used for background subtraction may be biased.

Analysis

The mean DC offset DN values will be binned on a 1 degree by 1 degree latitude-longitude grid based on the subsatellite position. A map will be constructed. Significant variations (TBD) over the map will indicate need for further analysis of variations correlated with orbital position.

The CCD must be powered-up with stowed elevation scan-mirror, and DN measurements recorded at low rate sampling intervals over several contiguous orbits. The set of orbits selected must be centered around the longitude of peak SAA activity (as determined from external data sets). The measurements are spatially segregated by latitude and longitude into four sets of time series as characterized by occurring within the SAA, within the two polar regions, and outside the previous three regions. The four sets of measurements are compared with a standard two-sample mean T-tests and two-sample variance F-tests in order to detect the influence of particle effects on the DC DN. If the tests indicate significant effects, the DC DNs associated with particle events provide prior information to the outlier-rejection algorithm.

Schedule

Monthly. More frequently when solar flare particles are known to have been trapped in the polar horns.

Presentation of Results

SAGE III internal memorandum.

Special Tools

PGS Toolkit, IMSL, IDL.

The time series analysis tools will be needed.

5.8. Coherent Noise in the Normalized Solar Scans

SAGE III data analyzed

Normalized solar scans from 150-300 km for all wavelengths in the CCD table.

Assumptions Tested

The SAGE III solar and lunar sampling was designed to minimize potential electronic noise crosstalk between data-take periods and other events occurring within the system (the ‘system’ in this context can also be taken to mean the S/C platform). Indicators of potential crosstalk are data sets whose sampling frequencies are greater than the scan mirror rotation rate.

Analysis

The CCD table wavelength solar scan data between 150 and 300 km are normalized relative to the mean scan. The set of normalized scans should represent a statistically stationary (mean and variance constant) time series of measurements. Spectral and cross spectral estimates are calculated, and statistically significant spectral and co-spectral peaks reported.

Schedule

Every solar event.

Presentation of Results

Frequencies corresponding to the significant power and cross spectral components reported to the SAGE III Algorithm Group. Significant components warrant special action in order to eliminate the contaminating noise from the atmospheric sampling.

Special Tools

PGS Toolkit, IMSL library, IDL.

5.9. Temperature Telemetry Change Point Detection

SAGE III Data Analyzed

Temperature telemetry low rate (0.1 Hz) sampling.

Assumptions Tested

During A&E the temperature telemetry will provide a ‘training’ baseline for predicting the excursions of the various measured temperatures within the instrument. Discontinuities in a given thermistor’s record may fall within the nominal limits of operation, however the location (in time) of these discontinuities may be rare or randomly distributed, and thus, not detectable by conventional spectral analysis techniques. Potential discontinuities may be correlated in time with geophysical sampling.

Analysis

An algorithm developed by Smid and Volf [*Smid and Volf*, 1996] forms the basis for detecting significant perturbations in each analog telemetry time series. The technique is based on spline interpolation and normalized Taylor coefficients, where the change point is detected as a singularity in the Taylor coefficients.

Schedule

Continuous process in SCF. Selected temperature telemetry time series must be made available to the SCF.

Presentation of Results

SAGE III memorandum. Significant times warrant special action to trace potential effects in the atmospheric sampling.

Special Tools

PGS Toolkit, IMSL library, IDL.

5.10. Auroral Emission in Lunar Measurements

SAGE III Data Analyzed

DN from lunar sampling in the high latitudes.

Assumptions Tested

The most intense auroral emissions in the high latitudes near peak solar activity originate from the interaction of solar electrons with atmospheric oxygen and nitrogen. The discrete spectral emissions occur in the visible in the wavelength range of the SAGE III spectrometer. The height of auroral emission occurs in the range from 80 to 130 km. The emissions are a radiant source detectable (auroral radiance TBD - but its probably close to lunar) by SAGE III, and hence a potential source of error in the lunar positioning and species retrieval.

Analysis

An auroral spectrum is used with the MLR technique to estimate the magnitude of auroral emission. Significant spectral components are written to an output file.

Schedule

The peak in solar activity for the SAGE III epoch will occur in 2002. Thus auroral activity might be countered immediately after the first SAGE III launch in 1998. For lunar tangent points poleward of 50° the lunar sampling auroral detection algorithm must be activated.

Presentation of Results

SAGE III memorandum.

Special Tools

IDL.

5.11. Lunar Disk Registration

SAGE III Data Analyzed

SAGE III DN during lunar calibration mode for the altitudes of 100 to 150 km.

Assumptions Tested

In order to retrieve extinction from lunar measurements, an accurate exoatmospheric lunar photometric function (analogous to the solar limb function) must be available. The lunar reflectance varies selenophysically as well as from S/C orientation during the event. The degree to which lunar disk scanning can be known is measured by comparison with lunar spectral reference maps derived from external sources.

Analysis

Lunar scans as a function of CCD table wavelength, time and elevation scan mirror position are compared with lunar disk maps for the same wavelengths and time. The RSS of the fit of the SAGE III single lunar scan to the expected scan derived from the lunar map is written to an output file.

Schedule

All lunar events.

Presentation of Results

SAGE III memorandum.

Special Tools

PGS Toolkit, IMSL library, IDL, and lunar spectral maps from Clementine and EOS.

5.12. Spatial Homogeneity Test

SAGE III Data Analyzed

SAGE III slant-path optical depth values sorted into altitude bins.

Assumptions Tested

Along-path spatial gradients can be determined from the multiple solar scans used to derive the transmission estimates constituting the L1B data product. As the instrument orbits, the sampled slant path corresponding to a given tangent altitude is not constant throughout an event. Thus, the variance in the data collected at that altitude will reflect the inhomogeneity in the atmosphere. A measure of inhomogeneity is passed to the L2 algorithm to act as an arbiter between several distinct branches of processing.

Analysis

Slant-path optical depth derived from solar scans as a function of CCD table wavelength, and elevation scan mirror position are binned into discrete altitude classes corresponding to refraction-corrected tangent altitudes in the atmosphere. Vertical gradients are removed, and the relative variance for the collection of binned data is calculated. The tangent-altitude dependent optical depth relative variances are written to an output file for each CCD table wavelength.

Schedule

All solar events.

Presentation of Results

SAGE III memorandum.

Special Tools

IDL.

5.13. Convergence Test

SAGE III Data Analyzed

SAGE III Pressure-Temperature (P/T) algorithm retrieval density.

Assumptions Tested

Within the SAGE III retrieval algorithm iterations are performed which are based on the value of the density differential between the current retrieved density and the previous density.

Analysis

Slant-path optical depth derived from solar scans as a function of CCD table wavelength, and elevation scan mirror position are binned into discrete altitude classes corresponding to refraction-corrected tangent altitudes in the atmosphere. Vertical gradients are removed, and the relative variance for the collection of binned data is calculated. The tangent-altitude dependent optical depth relative variances are written to an output file for each CCD table wavelength.

Schedule

All solar events.

Presentation of Results

SAGE III memorandum.

Special Tools

IDL.

5.14. MLR Wavelength Shift

SAGE III Data Analyzed

SAGE III L1B transmissions.

Assumptions Tested

The wavelength calibration performed outside the atmosphere (see section 5.3) is intended to provide the map between pixel number and wavelength relative to an assumed solar irradiance data set. More information is also available from the spectral information used by the PGE during gas species retrieval. Thus, the comparison between the wavelength calibration using the solar spectral database, and the spectral gas species data is an important measure of the performance of the wavelength calibration.

Analysis

The MLR technique is used to fit the pixel number to wavelength using the gas species spectra assumed in the PGE. The residuals of the MLR fit to these gas spectra are output.

Schedule

Selected gas retrievals.

Presentation of Results

SAGE III memorandum.

Special Tools

IMSL, IDL.

5.15. Dual Inversion Ozone Agreement

SAGE III Data Analyzed

SAGE III L1B transmissions.

Assumptions Tested

The ozone retrieval is performed using the MLR and standard LS techniques. The agreement (altitude dependent) between both retrievals is desirable in order that the best technique be selected for the appropriate range of altitudes.

Analysis

The ozone profile is retrieved both using the MLR technique and the conventional LS technique. The relative difference profile is calculated and written to output.

Schedule

All ozone profile retrievals.

Presentation of Results

SAGE III memorandum.

Special Tools

IMSL, IDL.

5.16. Cloud Product

SAGE III Data Analyzed

SAGE III extinction coefficients.

Assumptions Tested

The SAGE III Cloud Presence Data Products ATBD describes an extensive QA analysis. Quick look quality control will consist of data screening to see if the data conforms to certain expected criteria. Control studies, which still have to be specified, might include examination of the following:

- Do the extinction values and their ratios fall within specified ranges?
- Are the signal termination altitudes reasonable i.e., close to or below the tropopause, except for PSCs?
- Do profiles for consecutive events resemble one another in terms of their gross features? This particular quality control filter will need careful specification as cloud is not likely to be correlated.

Analysis

Full description appears in the Cloud Presence Data Products ATBD. Specific analyses will be specified in a later version of the QA Plan.

Schedule

See the Cloud Presence Data Products ATBD.

Presentation of Results

See the Cloud Presence Data Products ATBD.

Special Tools

IMSL, IDL.

6. Analysis Tools

6.1. SDP Toolkit

ECS provides a suite of routines for handling EOS-specific tasks. The Science Data Production (SDP) Toolkit is a fortran and C library useful for SAGE III file handling. Specific applications will be described in a future version of this document.

6.2. IDL

Much of the SCF QA processing requires use of Research Systems Incorporated Interactive Data Language (IDL). The system includes extensive support for data-handling, time series analysis, and high-end graphics. Specific applications will be described in a future version of this document.

6.3. IMSL

The ECS provides access to Visual Numerics IMSL libraries useful for time series analysis of certain SAGE III QA functions. Specific applications will be described in a future version of this document.